

~~CONFIDENTIAL~~
Security InformationMACHINE INDEXING PROJECTStatus and Plans

25X1A5A1

The previous work of this project has been directed to developing new methods for searching and correlating information. The development of new methods is necessary if modern high-speed electronic machines - in particular the Electronic Scanning System developed by IBM in collaboration with [] are to be used effectively by the Agency.

With modern electronic scanning equipment, the searching and correlation of information can be based on an entirely new principle. It is now possible for the searcher himself to set up at will new combinations of subject headings to control the machine's searching and selecting operations. This means, in practice, that the machine not only performs the laborious task of identifying documents of interest but in addition can accomplish this task in a more effective fashion than previously possible.

A new system of indexing is required to coordinate the analysis of documents with machine operations. The new indexing system is the key prerequisite to using the new scanning equipment effectively for searching and correlating information.

Previous work has developed basic methodology for constructing indexing systems for use with the new equipment. Such methodology might be compared to the plans for constructing a house. The materials of construction are the words and phrases - that is to say, the terminology - used to speak about the subject matter to be indexed. These materials of construction have been collected in the form of an extensive card file consisting of over 25,000 terms from all realms of science and technology. Preliminary analysis of these terms has been completed as to type of term (name of material, process, attribute, general concept) and field of use (general, chemistry, physics, biology, mechanics.) About 5,000 terms have been processed further - some 3,000 from biology to the point of final codification for machine indexing.

25X1A5A1

During the coming fiscal year, the work at [] will be directed to the construction of the new indexing system to cover the broad field of science and technology. In order to expedite this program and to ensure that its implications for statistical study of industrial activity are realized, it is planned to add a physicist and an economist to the [] group. 25X1

Since September 1, 1951 when this project was set up, attention has also been directed to incorporating the new methodology into the framework of Agency activities. A rudimentary indexing system based on the new methodology has been partly worked out and it is anticipated that this interim system will be ready for use on a trial basis before the end of 1952.

Tests in use of the interim system late in 1952 and early in 1953 will provide a basis for incorporating into Agency activities the more elaborate indexing system now in development at []

25X1A5A1

~~CONFIDENTIAL~~~~Security Information~~